Summer Research Workshops on Speech and Language: Progress and Plans (darpa98.tex)

Frederick Jelinek

Center for Language and Speech Research

Johns Hopkins University

Baltimore, MD, 21218

jelinek@jhu.edu

February 11, 1998

1 Vision of LVCSR Workshops

- 1993 & 94 held at CAIP - Rutgers (New Brunswick, NJ)

- 1995, 96, & 97 held at CLSP - Johns Hopkins (Baltimore, MD)

- Supported by DoD

- Time and place where exploration of innovations begins in an intensive manner

  - Testing of promising new approaches
  - Follow-up encouraged by government contracts
  - Student grants for follow-up

- Every year new projects

- A place where an ever widening international circle of researchers gets involved with improvements in LVCSR state-of-the-art.

- Initiation of students and their integration into the speech recognition community

2 The workshop environment at CLSP

- 6 weeks, starting in middle of July

- Work organized into projects with teams and team leaders

- Steering committee: Workshop director and coordinator, team leaders, govt. representative.

- Everyone gets workstation

- Each team assigned students who are familiar with software and computing environment

- Teams responsible for securing of software and data

- Outside lecturers speak on topics of interest to teams

- Teams report on their progress as it takes place

- The last two days of the workshop dedicated to reports on results

- Funds provided for selected student participants to continue research

  - Selection by an independent panel

- Participants are paid expenses; students and academicians get stipends

- Housing in apartments within walking distance from campus

3 Workshop planning process

- Process starts in organization meetings to which invited senior LVCSR researchers and potential participants

  - 1994 at Rutgers during Workshop and at later DoD LVCSR meeting
  - 1995 at conference center in Airlie, VA (2 day meeting)
  - 1996 at conference center in Leesburg, VA (2 day meeting)
  - 1997 at conference center in Airlie, VA (2 day meeting)

- Determine most burning LVCSR problems

  - Select those suitable as Workshop tasks

- Decide on team leaders

  - Form skeleton teams from those attending meeting

- Determine Workshop organization and needed resources

- Suggest names of potential participants

- Workshop coordinator named (1996: Lin Chase, CMU, 1997 & 1998: Joe Picone Miss.St)

  - Responsible for software and data provision

4   1997 LVCSR Workshop

- **Acoustic Processing**

  Team Leader: Andreas Andreou (CLSP/JHU)

  Members: Hynek Hermansky (OGI), Terri Kamm (DoD), Yasuhiro Minami (NTT), Christian Wellekens (Eurecom)

  Student members: Daniel Fain (CalTech), Juergen Luettin (IDIAP), Sarel van Vuuren (OGI)

  - **Aim:** Study and improve performance of specific components

- **Syllable – Based Speech Processing**

  Team Leader: George Doddington (DoD)

  Members: Andres Corrada (Dragon), Mark Ordowski (DoD), Joe Picone (Miss. State), Barbara Wheatley (DoD)

  Student members: Arvin Ganapathiraju (MissSt), Katrin Kirchoff (ICSI), Vaibhava Goel (CLSP/JHU)

  - **Aim:** Base acoustic model specification on syllabic strings
  - **Best result:** WER: $49.8 \implies 49.0$
  - **Tentative conclusion**: Results (comparison to monophone and triphone systems) seem to validate the hypothesis of integrity of syllables and relative independence from surroundings.

5 1997 LVCSR Workshop (continued)

- **Pronunciation Modeling**

  Team Leader: Michael Riley (AT&T)

  Members: Bill Byrne (CLSP/JHU), Michael Finke (CMU), Sanjeev Khudanpur (CLSP/JHU), John McDonough (BBN), Charles Wooters (DoD), George Zavaliagkos (BBN)

  Student members: Harriett Nock (Cambridge) Murat Saraclar (CLSP/JHU)

  - **Aim:**Replace static lexical baseforms underlying acoustic models by pronunciation networks reflecting the variety and dynamics of pronunciation found in conversational speech
  - **Reduction of WER**: $44.66 \Longrightarrow 41.4$

- **Discourse Language Modeling**

  Team Leader: Daniel Jurafsky (U. Colorado)

  Members: Marie Meteer (BBN), Klaus Ries (CMU), Liz Shriberg (SRI), Andreas Stolcke (SRI), Paul Taylor (Edinburgh), Carol Van-Ess Dykema (DoD)

  Student members: Noah Coccaro (UC Boulder), Becky Bates (BU), Rachel Martin (CLSP/JHU)

  - **Background:** Give and take of discourse seems capable of being cut up into coherent segments with characteristic modes of expression
  - **Aim:** Create a language model taking advantage of structure of discourse
  - **Observations:**
    * Bulk of error reduction is in relatively infrequent modes. Frequent modes are *answer* and *statement*
    * Utterance mode detection quite successful: error rate: 36.5 %
    * If utterance type were supplied by benevolent genie, WER $41.6 \Longrightarrow 40.8$
    * Best honest result: WER $41.6 \Longrightarrow 41.1$

6 1998 NSF Sponsored Speech and Language Workshop

- CORE NATURAL LANGUAGE PROCESSING TECHNOLOGY APPLICABLE TO MULTIPLE LANGUAGES:

    - *Team Leader*: Jan Hajic (Charles Univ., CR)

      Michael Collins (UPenn), Eric Brill (CLSP/JHU), Scott Miller (BBN), Doug Jones (DoD)

      Graduate Students: Barbora Hladka (CU, CR), Christoph Tillmann (Lehrstuhl, Aachen), Kyril Rybakov (CU, CR)

- RAPID SPEECH RECOGNIZER ADAPTATION FOR NEW SPEAKERS

    - *Team Leader*: Vassilis Digalakis (TUC)

      Enrico Bocchieri (AT&T), William Byrne (CLSP/JHU), Sanjeev Khudanpur (CLSP/JHU), Ashvin Kannan (Nuance), Ananth Sankar (SRI), DoD participant

      Graduate Students: Costas Boulis (TUC) John McDonough (JHU)

- DYNAMIC SEGMENTAL MODELS OF SPEECH COARTICULATION

    - *Team Leader*: John Bridle (Dragon UK)

      Li Deng (Waterloo), Joe Picone (MsState), Hywel Richards (Dragon UK), DoD participant

      Graduate Students: Jeff Ma (Waterloo) Mike Schuster (Nara, Japan)

- Targeted undergraduates: outstanding rising seniors who might not select language engineering for graduate study.

- Nationally advertised competition

    - Student requires faculty sponsor
    - Must provide evidence of proficiency in appropriate subjects
    - Short list of 25 out of 120 applicants selected at JHU
    - Teams will select 3 candidates out of short list
    - Workshop director selects the 2 participants

- Each student will have a mentor from his team who will introduce the subject to him

- Students treated as full members of teams: are invited to preliminary meetings.